R E P O R T R E S U M E S

ED 016 979 48 AL 001 158

DERIVATION OF PHONEME INVENTORIES BY NATIVE SPEAKER RESPONSES TO SYNTHETIC STIMULI. FINAL REPORT.

BY- SCHOLES, ROBERT J.
FLORIDA UNIV., GAINESVILLE

REPORT NUMBER BR-6-8486 PUB DATE 1 FEB 68
CONTRACT OEC-2-7-068486-2677
EDRS PRICE MF-$0.25 HC-$0.96 22P.

DESCRIPTORS- *PHONEMES, *PHONEMICS, *LANGUAGE RESEARCH, TAPE RECORDINGS, ARTIFICIAL SPEECH, CONSONANTS, PHONOLOGY, RECOGNITION, SPEECH SYNTHESIZERS, UNIVERSITY OF FLORIDA COMMUNICATION SCIENCES LABORATORY,

BECAUSE THE INTERFERENCE FROM HIS NATIVE LANGUAGE CAUSES A LINGUIST TO HEAR AND IDENTIFY THE SOUNDS OF A FOREIGN LANGUAGE IN TERMS OF HIS OWN, THE AUTHOR HAS PROPOSED A PROCEDURE DESIGNED TO (1) MAKE THE TASK OF PHONEMICIZING A LANGUAGE SHORT AND OBJECTIVE, (2) EQUATE THE PHONEMES OF A LANGUAGE WITH THE PERCEPTION OF THE USERS OF THAT LANGUAGE, AND (3) MAKE THE TASK OF TYPOLOGIZING LANGUAGES ON THE BASIS OF PHONEME PATTERNS OBJECTIVE. THIS PROCEDURE IS AS FOLLOWS--(1) TAPE RECORDINGS OF SYNTHETIC SPEECH SOUNDS ARE MADE. THESE TAPES CONTAIN A SUFFICIENT NUMBER AND RANGE OF STIMULI TO EXHAUST THE POSSIBLE PHONETIC BASES FOR THE PHONEMIC SYSTEMS OF THE LANGUAGES OF THE WORLD. (2) THESE TAPES ARE PRESENTED TO NATIVE SPEAKERS OF LANGUAGES UNDER INVESTIGATION. THE INFORMANT RESPONDS TO EACH STIMULUS BY SAYING WHETHER OR NOT IT SOUNDS LIKE ONE OF THE SPEECH SOUNDS OF HIS LANGUAGE, AND IF IT DOES, WHICH SOUND. (3) HIS RESPONSES ARE PLOTTED AGAINST ACOUSTIC MAPPINGS OF THE STIMULI TO DETERMINE THE NUMBER AND TYPES OF PHONEMES IN THE INFORMANT'S LANGUAGE. RESULTS SHOW THAT WHILE SOME OF THE TAPED SOUNDS HAVE TESTED SUCCESSFULLY (E.G. WEAK FRICATIVE STIMULI), OTHERS HAVE PROVED INADEQUATE (E.G., VOICELESS STOP STIMULI) AND REQUIRE FURTHER RESEARCH. (AMM)

DOCUMENT FILMED FROM BEST AVAILABLE COPY

ED016979

BR 6-8486-48
PA-48

FINAL REPORT

Grant No. OEC2-7-068486-2677

DERIVATION OF PHONEME INVENTORIES BY NATIVE SPEAKER RESPONSES TO SYNTHETIC STIMULI

February 1, 1968

U.S. EDUCATION OF HEALTH
EDUCATION AND WELFARE

Office of Education

Bureau of Research

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

AL 001 158

FINAL REPORT

Derivation of Phoneme Inventories by Native Speaker Responses to Synthetic Stimuli

Contract No. OEC2-7-068486-2677

Robert J. Scholes

February 1, 1968

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions do not, therefore, necessarily represent official Office of Education position or policy.

Communication Sciences Laboratory
University of Florida
Gainesville, Florida 32601

AL 001 158

Introduction

The purpose of this work is to develop a procedure which will:

1) make the process of phonemicization an objective task.
2) make the process of phonemicization a short task.
3) equate the phonemes of a language with the perception of the users of that language.
4) make the task of typologizing languages on the basis of phoneme patterns objective.

At present, phonemicizing a language is neither short nor objective. The standard procedure may be outlined as follows:

a) the analyst gathers a large corpus of utterances of the object language
b) the corpus is transcribed--i.e., written in some form of a phonetic notation
c) the transcribed corpus is scanned for cases of phonetically similar sounds in complementary distribution.

Although much work follows Step c), this will be sufficient to demonstrate our point.

The gathering and transcribing of the sufficiently large corpus typically takes several months--the analysis may occupy several more months. The procedure we will propose takes at most a few hours.

The biases of the analyst come in both the transcribing of the corpus and in its analysis. The phenomenon of hearing a foreign language as merely a distorted form of one's own is too well known to bear documentation here. Not only is the analyst's native language a cause of "interference" in his hearing of a new one, the same linguistic background will dictate in large part the set of pseudo-phonetic symbols with which he will transcribe it (e.g., note the Indo-European biases in the International Phonetic Alphabet).

The major deterents to objectivity and explicitness are, however, in Step c) and involve the notions "phonetically similar" and "complementary distribution". Not only has the obvious question as to just how similar two sounds have to be in order to be considered with respect to distribution has never been answered, the more basic question of what is meant by the term at all is largely unexplored. As for complementary distribution, any two sounds can be shown to be in complementary distribution if the context is extended far enough. For example, given a sound x and a sound y which are "phonetically similar" where x always occurs in the context A_B and y never occurs in the context A_B, we then say that x and y are allophones (variants) of a single phoneme; if, however, both x and y occur in the context X_B (where X may be anything, including A) we say that they are different phonemes since they contrast in the environment X_B. So long as the notion of context with respect to distribution is not defined, there is no way of determining which statement regarding x and y is true and which false.

Such problems do not arise at all in the procedure proposed herein.

The procedure we have in mind is as follows:

1. Tape recordings of synthetic speech sounds are made. These tapes contain a sufficient number and range of stimuli to exhaust the possible phonetic bases for the phonemic systems of the languages of the world.
2. These tapes are presented to native speakers of languages under investigation. The informant responds to each stimulus by saying whether or not it sounds like one of the speech sounds of his language, and if it does, which sound.
3. His responses are plotted against acoustic mappings of the stimuli to determine the number and types of phonemes in the informant's language.

Method

The tapes are first tested for adequacy in a forced-choice situation. Suppose, for example, a tape is prepared which varies the formant transitions between a rapid onset of voicing and an /a/-type vowel--thus exploring the cases of initial voiced stops. This tape would be presented to speakers of well-known languages who were furnished with varied inventories of responses; say, ba___, da___, ga___, none___, and who would be asked to respond to each stimulus in the appropriate way. If the results of these tests indicate informant categorizations compatible with recognized phonemicizations, the tape is accepted.

The second kind of testing involves open-ended responses of informants to these same stimuli. In these tests, no choices are furnished, and informants are asked to respond to stimuli with any words containing similar sounds or with no response at all.

The forced-choice testing determines the comparability of responses to the stimuli to known and accepted phonemicizations. Only after a tape has been judged to yield results compatible with known facts regarding well-analyzed languages may it be used in open-ended testing.

The purpose of the open-ended testing, is, of course, the phonemicization itself.

Results

During the term of this contract, numerous tapes of synthetic speech sounds have been prepared and tested. Some of these have been rejected as inadequate, others have been modified, others retained as generated. We will describe results of our work in terms of those tapes and stimuli with which we are working at this point. There is no assurance, at this time, that these will prove to be adequate for our purposes; they may well require modifications in terms of some or all of the acoustic parameters involved.

At present, we are working with three sets of stimuli: voiceless stops, voiceless strong fricative, and voiced stops. The stop consonant stimuli were done on the Haskins Laboratories speech synthesizer; the fricatives were done with equipment in the Communication Sciences Laboratories. The generation and specification of these stimuli will be described in detail in the appendix; here we will discuss the generalities of their production and perception.

Voiceless Strong Fricatives

Voiceless strong fricatives are those fricatives which do not require formant transitions into contiguous vowels for their identification--e.g., /s/, /š/. Such sounds were generated by driving a formant vowel synthesizer with a noise source and increasing tape speeds to gain higher frequencies of the formants. Trained phoneticians were asked to judge the place of articulation of these stimuli and it was found that:

1) there was some patterning in their judgments in that the highest formant frequencies were judged to be /f/-like, next highest were /s/-like, next highest were /ʃ/-like, and the lowest were /h/-like.
2) most stimuli were judged as "none", meaning that the listener could not assign a place of articulation.

Voiceless Stops

We are currently working with a set of 62 synthetic voiceless stops in which all parameters are held constant throughout the set save the second and third formant transitions over the first 50 msecs. F2 initial values go from 543 cps to 2837 cps in ten steps, terminating at 1232 cps; F3 initial values go from 1190 to 3530 cps in seven steps, each transition terminating at 2525 cps.

These stimuli have been presented to speakers of various languages who were asked to respond with example words to each of the stop +/a/ stimuli they found to resemble sounds of their languages; they have been presented to English speakers in both forced-choice and open ended tasks. In general, each of these studies has shown the stimuli to be quite adequate for pre-velar places of articulation (/p/, /t/), but questionable for velar and post-velar stops (/k/, /q/). There is, therefore, little more to be said until the questions of the acoustic cues relevant to the production and perception of back stops has been clarified. A figure is given in the appendix which illustrates the forced-choice p-t-k-none judgments of 25 English speakers for these stimuli.

Voiced Stops

The initial voiced stops seem, according to the testing done to date, to be satisfactory. The identification by 21 English speakers shows sufficient numbers of b, d, and g judgments, and these judgments appear in the expected order. The stops differ from the voiceless set with respect to two parameters--the type of source over the first 35 msecs, and the amplitude of the fricative over the first 25 msecs.

Discussion

The results of these investigations have been less than conclusive due to a number of factors. First among these is the lack of information on the acoustic cues for post-velar speech sounds. Synthesizing a set of consonants is a reasonably straight-forward task so long as one deals with places of articulation within the p-t-k, b-d-g, range; if, however, one wants to exhaust the possible distinctions within the entire bilabial to glottal spectrum, then a good deal of basic analysis on post-velars must be done. We have attempted to circumvent this analysis by simply extending the known parameters (supplied, for the most part , by researchers at Haskins) for front-to-back distinctions in our synthesis. This has not been shown to be adequate, although it may well turn out to be so later. We have to some extent gotten /k/-/q/ type judgments, but not nearly as many as we would like, and those stimuli judged to be post-velar on some occasions do not appear to be clearly or consistently so judged.

The lack of post-velar judgments is also a factor of the language backgrounds of the subjects used. We have not found enough native speakers of languages which have phonemic post-velars to be able to say with much confidence what their judgments will typically be. This lack will be made up in further testing.

At this point, therefore, our attention for all stimuli-stops and fricatives-is being devoted to the adequacy of the parametric values for velar and post-velar places of articulation. These studies involve tests for the identification and discrimination of the stimuli by speakers of various languages.

Conclusions

For the voiceless stop stimuli, we appear to have inadequate tapes for the purposes described. Problems in these stimuli are that 1) too many of the stimuli are of the bilabial and alveolar types, while 2) not enough of them are of the velar and post-velar types. Research on these stimuli will be continued in an attempt to identify and correct the reasons for these problems.

For fricatives, it was found necessary to distinguish between weak and strong types. Weak fricative stimuli have been prepared and tested with some indications of success. In these stimuli formant transitions like those used for stops are employed over a longer time span than in the stops and with less abrupt onsets. Strong voiceless fricatives were generated with a shaped noise source. This stimuli have been judged to be speech-like by English speakers.

The basic question to be asked of all of these stimuli cannot, however, be answered at the moment. Whether they are adequate for the identification of such speech sounds in the languages of the world has not been established. It would appear that they are not to the extent that stops and fricatives of post-velar places of articulation are not exhaustively illustrated in the stimuli. This shortcoming will be rectified in the future research.

Summary

Having shown that synthesized vowels could be used in open-ended identification tasks to establish the vowel phoneme patterns of most languages the techniques used in this work are being applied to consonantal stimuli.

Four sets of consonantal stimuli have been selected for this work:

1) voiceless stops before /a/
2) voiced stops before /a/
3) voiceless weak fricatives before /a/
4) voiceless strong fricatives, isolated

Attempts have been made to generate sets of these stimuli, using equipment here at CSL and at Haskins Laboratories, such that the range and incremental differences within each such set is sufficient to cover the known place-of-articulation types within each manner described. Each stimuli tape should then as a whole be easily identifiable according to manner, and within each such tape there should be stimuli illustrative of every known type--from bilabial to glottal (insofar as such places of articulation are known to be employed in natural languages).

Success to date has been characterized by various degrees for each set of stimuli. The voiced stop onsets appear to be satisfactory at least for the b-d-g- range, but no data on their utility for languages employing finer or more far ranging divisions have been gathered. The voiceless stops are not identified as velar or post velar, even by English speakers; these stimuli require at least moderate revision. The shaped-noise stimuli used for strong fricatives can be identified as /s/ and /ʃ/ (but also /f/ and /h/) by phoneticians, but their lack of naturalness has hindered further testing. The voiceless weak fricatives before /a/ have defied our attempts altogether. Two runs have produced stimuli which did not sound different from each other at all and stimuli which were much to loud with respect to the following vowel to be natural.

Consequently, the prognosis for this research centers on the generation of adequate synthetic stimuli, rather than on extensive testing. At worst, some information on what not to do has been gathered; at best, the acoustic characteristics of these consonant types is becoming better understood.

APPENDIX

Details of Synthesis and Test Results

A Descriptive Chronology of Research Activities

February 1, 1967-April 31, 1967

Stimuli delimited to
- Voicing in stops
- Voicing in fricatives
- Place in weak fricatives
- Place in strong fricatives
- Place in voiceless stops
- Place in voiced stops

Exploratory visit to Haskins (April 10-14)
Dr. E. C. Trager visits CSL

May 1, 1967-July 31, 1967

Preparation of Stimuli

Voiceless stops--place range--75 stimuli
variations in F1, F2, F3 transitions

F1 = 260, 412, 562 cps
F2 = 543, 769, 1075, 1620, 2234, 2837 cps
F3 = 1190, 1524, 1849, 2525, 3195, 3530 cps
steady state values

Fo	120 cps
F1	743 cps
F2	1075 cps
F3	2525 cps
F1 amp	0 db
F2 amp	-3 db
F3 amp	-9 db
Fric amp	n/a
Overall amp	0 db
Source	buzz
Type of Fric.	n/a

Voiceless weak fricatives--place--67 stimuli

All parameters like those above except:
no F3 at 1524
onset time is 75 (vs 50) msec.

Voiceless strong fricatives--place--96 stimuli

These stimuli were generated by driving the CSL formant synthesizer with a noise source and raising frequency by mulitiplying tape speed.

Visit to Haskins (2) June 15-17, July 20-24

August 1, 1968-October 31, 1967

Testing the stimuli

Voiceless strong fricatives (shaped noise)

288 shaped noises were judged according to place of articulation by four trained phoneticians. Only 34 of these received same-place judgments by three of the judges--the others being assigned to two or more places by the judges. In general, those shaped noises and corresponding judgments which seemed clear were:

F1 = 400 cps
F2 = 900 - 1200 cps /h/

F1 = 800 - 1200 cps
F2 = 1800 - 2400 cps /ʃ/
H-P filt 1 kc

F1 = 2400 - 3200 cps
F2 = 3600 -4800 cps /s/
H-P filt 2 kc

F1 = 4000 - 6400 cps
F2 = 7200 - 9600 cps
H-P filt 2 or 3 kc /f/

Voiceless weak fricatives before /a/

Judgments indicated that differences in fricative onsets were not detectable, and plans were made to regenerate the stimuli.

Voiceless stops before /a/

Results from seven trained phoneticians who judged place-of-articulation indicated that F1 initial values did not contribute to such judgments. Both F2 and F3 transitions do contribute to such judgments. /p/ judgments were most frequent when both F2 and F3 initial values are low; /t/ judgments occur when F2 and F3 initial values are both high; /k/ judgments occur when F2 transition is slightly rising and F3 transition is falling.

November 1, 1967-January 31, 1968

Generating Stimuli

Voiceless stops before /a/

A new set of initial voiceless stops was generated which extended the range of F2 and F3 transitions and used only F1 transition. This new set contains 62 stimuli where the initial values for F2 and F3 were:

F2 = 543, 769, 996, 1232, 1465, 1695, 1920,
2156, 2387, 2615, 2837 cps
F3 = 1190, 1524, 1849, 2180, 2525, 2862
3195, 3530 cps

Steady state values: F2 = 1232
F3 = 2525

Voiced stops before /a/

A set of 62 initial voiced stops were generated which are exactly like the voiceless stop onsets except for the onset of voicing (buzz source) which here occurs during the very first time-slot. (for the voiceless stops, the first 35 msecs have a hiss source).

Voiceless weak fricatives before /a/

An attempt was made to regenerate clear voiceless weak fricatives. In this set, the fricatives are better heard than in the first, but do not sound "natural"--they are too strong, and resemble stop onsets as much as weak fricative onsets.

Visit to Haskins was made November 1-5.

Testing Stimuli

Voiceless stops

The voiceless stops have been tested in a variety of ways. They have been presented to speakers of English and other languages who have been asked to classify them in both forced-choice and open-ended tasks.

Twenty-five speakers of various languages were run in an open-ended classification test (give an example word for sounds like those in your language)

Twenty-five speakers of English were run in a forced-choice test (is the sound p, t, k, or none of these).

Both types of tests showed a serious lack of velar and post velar initial stops in the set of stimuli, but quite clear /p/ and /t/ types.

Voiced stops

Voiced stop stimuli are being run in forced-choice tests (b, d, g or none) with 25 speakers of English. Results indicate that this set of stimuli may be satisfactory.

Voiceless weak fricatives

These stimuli, although not adequate for their intended use are being used by the principal investigator and by a doctoral candidate in laterality experiments.

Projected Activities

Voiced and voiceless initial stops

These stimuli will be regenerated where, according to results to date, the range of values used in F2 and F3 transitions need not be so large, the number of such transitions within the lesser range can be expanded (i.e., smaller increments in initial values). Pending the outcome of the results of testing the voiced stop onsets, it would appear that the consonantal onset duration should be extended from the present 50 msecs to 75 or so. It would also seem that the steady state values of the three formants should be altered so as to permit perception of the onsets to be interms of being before an unrounded vowel--the present patterns correspond to results of the perception of stops before rounded vowels--i.e., a p-k-t (rather than p-t-k) progression.

Fricatives

Strong fricatives will be attempted by modifying the existing tapes of shaped noise stimuli and by programming the Haskins synthesizer.

Weak fricatives will be re-done on the Haskins synthesizer attempting to reach a happy medium between the present tapes where they are either too weak or too strong in relation to the following vowel.

The Voiceless Stop Stimuli, Set 1

The voiceless stop stimuli were generated from a basic pattern on the Haskins Laboratories speech synthesizer. On this machine one specifies each of eleven acoustic parameters for each of n time slots where the time slots may be 5 or 10 msecs. in length. See Figure 1 for parametric values.

We used 65, five msec., time slots for these stimuli, giving an overall length of 325 msecs. The initial 50 msecs is termed the onset and all variations in the set of stops occur in this segment. The middle 250 msecs is termed the steady state and no variations occur in this segment; the only parameter which changes at all is the fundamental frequency which rises and falls to give a normal contour. The final 75 msecs. is termed the decay and no variations within the set occur here; the decay segment is that portion wherein the amplitude and fundamental frequency are lowered to give a natural sounding to the syllable.

The entire syllable thus consists of a voiceless stop onset plus an /a/ vowel steady state and decay. The steady state and decay parameter values for the eleven parameters are given below:

Parameter	Acoustic Analogue	Value Used
0	F1 frequency	743 cps
1	F2 frequency	1075 cps
2	F3 frequency	2525 cps
3	F1 amplitude	0 db
4	F2 amplitude	-3 db
5	F3 amplitude	-9 db
6	Fricative amplitude	none
7	Fundamental frequency	peak at 120 cps
8	Over-all amplitude	peak at 0 db
9	Type of source	buzz
10	Type of fricative	none

The variations in the set of stimuli were located in the frequency onsets of the first, second, and third formants, all other parameters being constant throughout the set. Each formant frequency was made to progress linearly from an initial value to the steady state value over the period of 50 msecs. The voiceless stop nature of the onset is given by an initial 5 msec. burst which has a noise source driven through the formants at considerable amplitude; this is followed by a 35 msec. portion where the noise source drives the output through the formants at a reduced amplitude; gradually the buzz (voicing) source takes over, coming on full at 45 msec.

The stimuli generated, with the initial formant values for each of the three formants is given below:

			F3				
F1	F2	1190	1524	1849	2525	3195	3530
	543	1	2	3	4	5	6
	769	7	8	9	10	11	12
260 X	1075	13	14	15	16	17	18
	1620			19	20	21	22
	2234				23	24	25
	2837					26	27
	543	28	29	30	31	32	33
	769	34	35	36	37	38	39

412	X	1075	40	41	42	43	44	45
		1620			46	47	48	49
		2234				50	51	52
		2837					53	54
		543						
		769	55	56	57	58	59	60
562	X	1075	61	62	63	64	65	66
		1620			67	68	69	70
		2234				71	72	73
		2837					74	75

The gaps in the chart above are due to the fact that no initial formant value is used which is higher than the initial value of the next-higher formant--to do so would produce an unnatural configuration. See Figures 2 and 3 for results.

Weak Voiceless Fricatives, Run 1.

The weak voiceless fricatives were produced in a similar way to the stops: Dr. Haggard generated an /fa/ pattern on which variations in the formant frequency initial values were made. We are thus looking at the affect of first, second, and third formant transitions on the place perception of voiceless weak fricatives. The stimuli by F1, F2, and F3 initial values used in this tape were as below:

F1		F2	1190	1849	2525	3195	3698
			1				
		543	1	2	3	4	5
		769	6	7	8	9	10
260	X	1075	11	12	13	14	15
		1620		16	17	18	19
		2234			20	21	22
		2837				23	24
		543	25	26	27	28	29
		769	30	31	32	33	34
412	X	1075	35	36	37	38	39
		1620		40	41	42	43
		2234			44	45	46
		2837				47	48
		543					
		769	49	50	51	52	53
562	X	1075	54	55	56	57	58
		1620		59	60	61	62
		2234			63	64	65
		2837				66	67

The same restriction on formant crossings applies here as to the voiceless stops.

Noise-driven vowel synthesizer

The vowel synthesizer in use here at the Communication Sciences Laboratory was driven by a noise source to produce noise spectra with the following first and second formant values:

F1: 300, 400, 500, 600, 700, 800 cps
F2: 600, 800, 1000, 1200, 1400, 1600 cps

The resulting stimuli were then boosted up in frequency by increasing tape speeds by factors of 2, 4, and 8. See Table 1 for results.

Stimuli. Run 2.

The voiceless stop plus /a/ syllables used in these studies were generated on the terminal analog speech synthesizer of Haskins Laboratories. This synthesis system permits the user to specify each of eleven acoustic parameters of the signal at every five milliseconds. Once a signal has been constructed, changes by parameter or by time slot can be made by very simple instructions.

For these stimuli, 61 five msec. time slots were used, giving the syllables an overall length of 305 msecs. Of this total length, the first 50 msecs are of interest in the studies to be described. The other 255 msecs contain no parametric changes other than a linear drop in fundamental frequency over the final 16 time slots (80 msecs). The fundamental frequency thus rises from 116 cps at time slot to 138 cps at time slot 17, remains at 138 cps until it begins a gradual fall at time slot 45. Parameter 11, which is concerned with the type of fricative -/s/, /ʃ/, or /f-θ /- is not used at all for these stimuli.

During the first 50 msecs various parameters are specified so as to produce the general class of initial voiceless stops. These parameters and the changes specified are:

Para 10: Type of source. A hiss (noise) source is specified through the first 30 msecs, buzz (voice) source is specified at time slot 7 (35 msecs) and continued throughout the syllable.

Para 9: Overall amplitude. Overall amplitude rises steeply from an initial value of -10.5 db to full volume at time slot 6 (30 msecs).

Para 6: Fricative amplitude. The fricative source is given an amplitude of -3 db through the first 15 msecs., then brought down to zero at 25 msecs.

Para 5: F3 amplitude. The amplitude of F3 begins at -15 db, falls to -21 db, then rised gradually to its steady state value of -12 db.

Para 4: F2 amplitude. The amplitude of F2 begins at a value of -6 db, falls immediately to -15 db and rises gradually to its steady state value of -3 db.

Para 3: F1 amplitude. The amplitude of F1 begins at -9 db and rises gradually to its steady state value of -0 db.

Para 3 and Para 1, the frequencies of the third and second formants will be discussed separately since these are the variable across stimuli.

Para Ø: F1 frequency: The first formant transition is a linear rise from 260 to 743 cps over the first 50 msecs.

The second and third formant values over the steady state portion of the stimuli are F2 = 1232 cps and F3 = 2525 cps. These values, together with the F1 value of 743 cps give a slightly fronted, slightly rounded (due to energy of F3) vowel /a/. These formant value are quite compatible with the attempt to use a vowel which is found throughout the languages of the world (see Scholes, 1966).

The sixty-two stimuli used in these experiments were constructed by specifying 11 different initial frequency values for F2 and 8 different initial frequency values for F3. The transitions from these initial values were linear and occurred during the first 50 msecs. of the stimulus. One of 11 F2 initial values--543, 769, 996, 1232, 1465, 1695, 1920, 2156, 2387, 2615, or 2837 cps was combined with one of 8 F3 values--1190, 1524, 1849, 2180, 2525, 2862, 3195, or 3530 to produce each stimulus. There was the restriction that F2 could not be higher than F3; a stimulus with an initial value of F2 at 2837 cps and F3 at 2180 cps, for example, was not generated.

The various transitions of F3 and F2 are shown in Figure 2. The left-most column shows the counts in cycle-per-second values; and the third column is the magnitude of change, in cycles-per-second, between the initial and steady state values of the formant frequencies--a higher initial value is designated by a plus, a lower value by a minus. These third column figures will be used to designate stimuli throughout the remainder of this report.

Table 2 shows the forced-choice identifications of the voiceless stop stimuli. Table 3 shows the forced-choice identifications of the voiced stop stimuli.

Synthesis Parameters	
Type of Fricative	Three settings corresponding to /f, θ/, /s/, or /S/
Type of Source	Buzz, Hiss, or Buzz + Hiss
Overall Amplitude	16 increments from -a to 0 db in 1.5 db steps
Fundamental Frequency	64 increments from 70 to 196 cps in 2 cps steps
Fricative Amplitude	7 increments from -a to 0 db in 3 db steps
Third Formant Amplitude	16 increments from -a to 0 db in 3 db steps
Second Formant Amplitude	"
First Formant Amplitude	"
Third Formant Frequency	16 increments from 1190 to 3640 cps in 150 cps steps
Second Formant Frequency	32 increments from 543 to 2910 cps in 80 cps steps
First Formant Frequency	32 increments from 100 to 804 cps in 27 cps steps

1 2 3 4 5 6 7 8 9 1 1 1 1 1 1 1 1 1 1 2 ... 5 5 5 5 5 5 5 5 5 6 6
0 1 2 3 4 5 6 7 8 9 0 0 1 2 3 4 5 6 7 8 9 0 1

Five msec time steps

Figure 1. The Haskins Synthesis Parameters

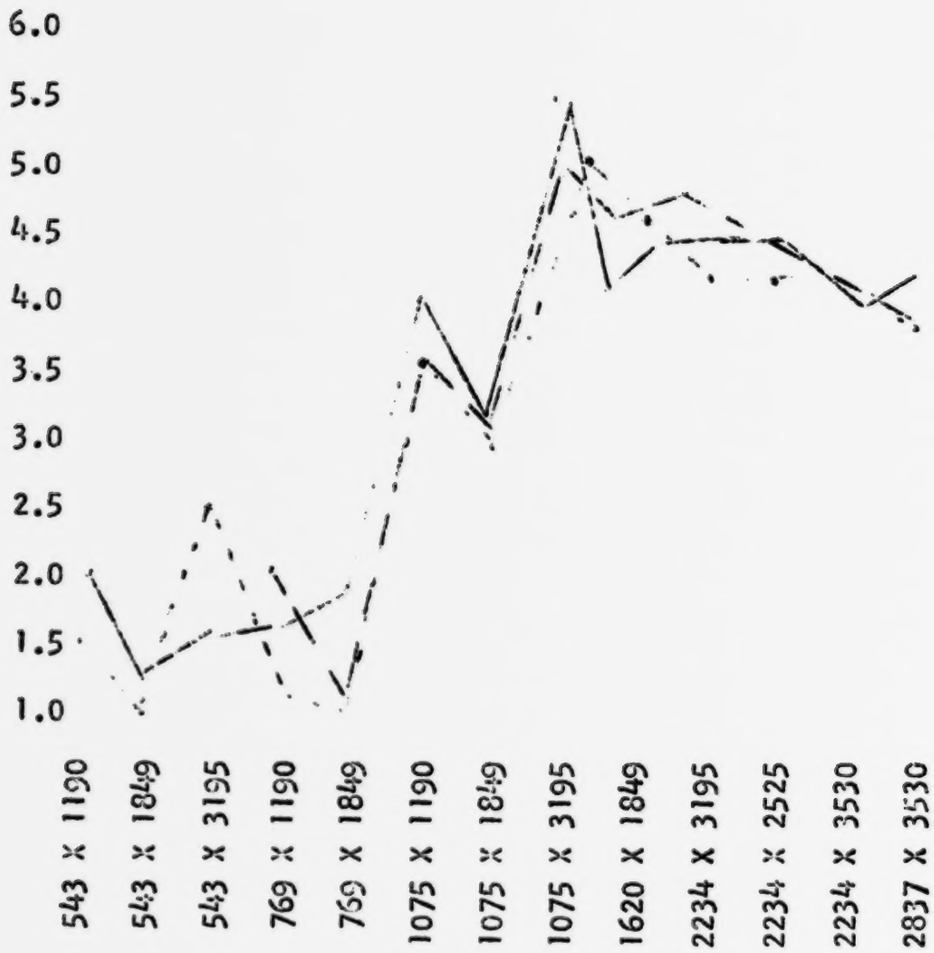

Figure 2. Effect of F1 transition on the perception of place. Not all values shown. Solid line = F1 at 260 cps.; dash line = F1 at 412 cps.; broken line = F1 at 562 cps.

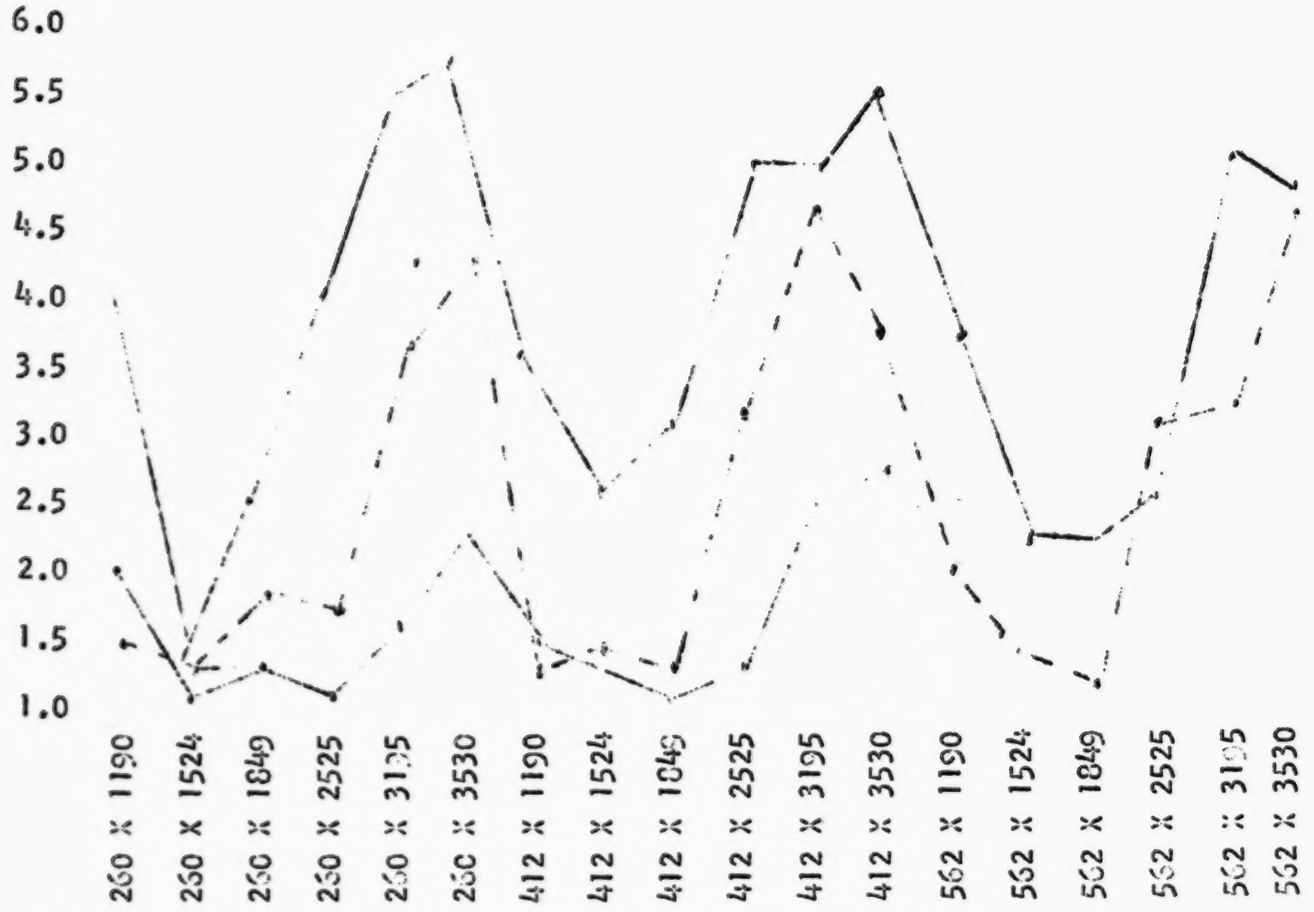

Figure 3. Effect of F2 transition on the perception of place. Lower solid line = F2 at 543 cps.; dotted line = F2 at 760 cps.; upper solid line = F2 at 1075 cps.

	1	2	3
300 × 900			
300 × 1000			
300 × 1100			
300 × 1200			
400 × 900			
400 × 1000			
400 × 1100			
400 × 1200			
500 × 900			
500 × 1000			
500 × 1100			
500 × 1200			
600 × 900			
600 × 1000			
600 × 1100			
600 × 1200			
700 × 900			
700 × 1000			
700 × 1100			
700 × 1200			
800 × 900			
800 × 1000			
800 × 1100			
800 × 1200			

	1	2	3
600 × 1800			
600 × 2000			
600 × 2200			
600 × 2400			
800 × 1800			
800 × 2000			
800 × 2200			
800 × 2400			
1000 × 1800			
1000 × 2000			
1000 × 2200			
1000 × 2400			
1200 × 1800			
1200 × 2000			
1200 × 2200			
1200 × 2400			
1400 × 1800			
1400 × 2000			
1400 × 2200			
1400 × 2400			
1600 × 1800			
1600 × 2000			
1600 × 2200			
1600 × 2400			

	1	2	3
1200 × 3600			
1200 × 4000			
1200 × 4400			
1200 × 4800			
1600 × 3600			
1600 × 4000			
1600 × 4400			
1600 × 4800			
2000 × 3600			
2000 × 4000			
2000 × 4400			
2000 × 4800			
2400 × 3600			
2400 × 4000			
2400 × 4400			
2400 × 4800			
2800 × 3600			
2800 × 4000			
2800 × 4400			
2800 × 4800			
3200 × 3600			
3200 × 4000			
3200 × 4400			
3200 × 4800			

	1	2	3
2400 × 7200			
2400 × 8000			
2400 × 8800			
2400 × 9600			
3200 × 7200			
3200 × 8000			
3200 × 8800			
3200 × 9600			
4000 × 7200			
4000 × 8000			
4000 × 8800			
4000 × 9600			
4800 × 7200			
4800 × 8000			
4800 × 8800			
4800 × 9600			
5600 × 7200			
5600 × 8000			
5600 × 8800			
5600 × 9600			
6400 × 7200			
6400 × 8000			
6400 × 8800			
6400 × 9600			

Table 1. Specification of filtered, shaped noises

Third Formant Initial Values

1190	1524	1849	2180	2525	2862	3195	3530	Second Formant Initial Values
p	p	p	p	p	p	p	p	543
p	p	p	p	p	p	p	p	769
p	p	p	p	p	p	p	p	936
	p	p	p	p	p	p	p	1232
	p	p	p	p	p	p	p	1465
		-	k	t	t	t	t	1695
			t	t	t	t	t	1920
				t	t	t	t	2156
				t	t	t	t	2387
					t	t	t	2615

Table 2. Preferred choices in forced-choice test. 25 subjects

Voiceless initial stops, Run 2.

Third Formant Inital Value

1190	1524	1849	2180	2525	2862	3195	3530	Second Formant Initial Values
b	b	b	b	b	b	-	n	543
b	b	b	b	b	b	-	n	769
b	b	b	b	b	b	b	-	996
	b	b	b	b	b	-	g	1232
	g	g	g	g	d	g	d	1465
		g	g	-	g	d	-	1695
			g	g	d	g	-	1920
				g	d	d	g	2156
				g	g	g	g	2387
					g	g	g	2615
						g	g	2837

Table 3. Majority judgments for voiced stop = /a/ stimuli. Forced-choice; p-d-g-none

N = 21 English speakers

- = no majority opinion